Context-free languages and primitive words

Pál Dömösi *

Combinatorial properties of words play an important role in mathematics and theoretical computer science. One of the well-known open problems is related to the language of primitive words. A word is called *primitive* if it is not a repetition of another word. (Thus the empty word is non-primitive.)

We conjectured that the language Q of all primitive words over a non-singleton alphabet is not context-free (Dömösi, S. Horváth, M. Ito [1991]). The problem seems to be simple but we could not solve it yet.

Apart from the conditions of Wise Lemma (D. S. Wise [1976]), Q has all well-known iteration conditions of context-free languages (P. Dömösi, S. Horváth, M. Ito, L. Kászonyi, M. Katsura [1992,1993]). [1] Another test of context-freeness is the so-called Interchanging Lemma (W. Ogden, R. J. Ross, K. Winklmann [1982]). It is also proved that Q fulfils the conditions of this test (S. Horváth [1995]). Therefore, Q resists almost all well-known tests of context-freeness.

It is also well-known that an intersection of a regular and a context-free language is again a context-free language. Therefore, if we find a regular language R such that $R \cap Q$ is not context-free then we can show that Q is not context-free. By results of L. Kászonyi and M. Katsura [1996, 1997, 1999a, 1999b], this direction also seems to be hopeless.

Maybe an appropriate homomorphic characterization of languages could help to prove our conjecture about the context-freeness of Q. (N. Chomsky and M. P. Schützenberger [1963], R. J. Stanley [1965]), S. Hirose and M. Yoneda [1985], P. Dömösi and S. Okawa [2003]).

*Nyíregyházi Főiskola, Matematika és Informatika Intézet,
H-4400 Nyíregyháza, Sóstói út 31/B, e-mail: domosi@nyf.hu

[1] Note that the applicability problem of the Wise Lemma is equivalent to the original problem.)

1 Concepts

alphabet - non-empty and finite set - Σ

non-trivial alphabet - $|\Sigma| > 1$

letter - an element of the alphabet

word - a finite string of letters

$p = x_1 \cdots x_k, q = x_{k+1} \cdots x_\ell, x_1, \ldots, x_\ell \in \Sigma \Rightarrow pq = x_1 \cdots x_\ell$
$p^0 = \lambda$, where λ is the empty word; $p^k = pp^{k-1}; p^* = \{p^k \mid k \geq 0\}; p^+ = p^* \setminus \{\lambda\}$

the set of all words : Σ^* including the empty word λ

the set of non-empty words : $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$

primitive word: it is not a repetition of another word

language : $L \subseteq \Sigma^*$

bounded language : there exists $w_1, \ldots, w_n \in \Sigma^*$ with $L \subseteq w_1^* \cdots w_n^*$

slender language: $\exists c > 0 : \forall n > 0 : |L \cap \{w \in \Sigma^* : |w| = n\}| \leq c$

k-slender language: $\exists c > 0 : \forall n > 0 : |L \cap \{w \in \Sigma^* : |w| = n\}| \leq cn^k$

paired loop language : $\{uv^n wx^n y : u, v, w, x, y \in \Sigma^*, n \geq 0\}$

Non-crossing 1-time paired loop language : paired loop language

Non-crossing $k+1$-times paired loop language : $k \geq 1$: $L = \{uv^n Lx^n y | n \geq 0\}$ for some $u, v, w, x, y \in \Sigma^*$ and a non–crossing k-times paired loop language
Moreover, if L_1 is a non-crossing k-times paired loop language and L_2 is a non–crossing ℓ-times paired loop language then $L_1 L_2$ is called a non-crossing $k + \ell$-times paired loop language

Non-crossing paired loop language : non-crossing k-times paired loop language for some k

All primitive words over Σ : Q

grammar - $G = (V, \Sigma, S, P)$, where

$$P \subset \{W \to Z \mid W \in (V \cup \Sigma)^*, Z \in (V \cup \Sigma)^+\}$$

direct derivation : $W \underset{G}{\Rightarrow} Z$,
where $W = W_1 W' W_2, Z = W_1 Z' W_2, W' \to Z' \in P$

derivation : $W \overset{*}{\underset{G}{\Rightarrow}} Z$, where either $W = Z$ (and then $n = 1$) or
$\exists W_1, \ldots, W_n : W_1 = W, W_n = Z, W_i \underset{G}{\Rightarrow} W_{i+1}, i = 1, \ldots, n-1$ (and then $n > 1$)

language generated by G : $L(G) = \{p \in \Sigma^* \mid S \overset{*}{\underset{G}{\Rightarrow}} p\}$

The grammar $G = (V, \Sigma, S, P)$-t is called *i-type* if
- $i = 0$: and there exists no further restriction *(Phrase-structural grammar)*
- $i = 1$: All elements of P has the form $W_1 Q W_2 \to W_1 R W_2$, where $W_1, W_2 \in (V \cup \Sigma)^*$, $Q \in V$, $R \in (V \cup \Sigma)^* \setminus \{\lambda\}$; or $S \to \lambda \in P$, but then S does not appear on the right side of any derivation rule *(Context-free grammar)*
- $i = 2$: All elements of P have the form $Q \to R$, where $Q \in V_N$, $R \in (V_N \cup V_T)^*$. *(Context-free grammar)*
- $i = 3$: All elements of P have one of the forms $Q \to pR$ or $Q \to p$, where $Q, R \in V_N$, $p \in V_T^*$. *(Right-linear grammar)*

Conjecture (Dömösi-Horváth-Ito, 1991) : The language of all primitive words over a non-trivial alphabet is not context-free.

1.1 Iteration Properties

Y. Bar-Hillel, M. Perles and S. Shamir [1961]:

Theorem 1 [*Bar-Hillel Lemma*] *For each context-free language L there exists a positive integer n with the following property: each word z in L, $|z| > n$, is of the form $uvwxy$, where $|vwx| \leq n$, $|vx| > 0$, and uv^iwx^iy is in L, for all $i \geq 0$.*

We say that L satisfies the Bar-Hillel condition if it has the above properties in Theorem 1.

G. Borwein (published by C. M. Reis and H. J. Shyr [1978]):

Theorem 2 [*Borwein Lemma*] *Let $u \in \Sigma^+, u \notin a^+, a \in \Sigma$. Then at least one of ua, u must be primitive.*

P. Dömösi, S. Horváth, M. Ito, and L. Kászonyi [1994]:

Theorem 3 *Q satisfies the Bar-Hillel condition.*

Proof: Let $w \in Q, |w| \geq 2$ and denote by m the maximal positive integer with $w = pa^mq, p, q \in \Sigma^*, a \in \Sigma$. By the well-known Borwein Lemma (see Theorem 2), if $b \in \Sigma, pq \in \Sigma^+ \setminus b^+$ and $pbq \in \Sigma^+ \setminus Q$ then $pq \in Q$. If $m > 1$ then, using this result, either $pa^{m-1}q \in Q$ or $pa^{m-2}q \in Q$. On the other hand, by the maximality of m we have $pa^{m+j}q \in Q, j \geq 0$. Therefore, we obtain $uv^iwx^iy \in Q, i \geq 0$ if either $pa^{m-1}q \in Q, u = pa^{m-1}, v = a, w = x = \lambda, y = q$, or $pa^{m-2}q \in Q, u = pa^{m-2}, v = a^2, w = x = \lambda, y = q$.

If $m = 1$ then $pq \in Q$ (by $|pq| > 0$) and $pa^jq \in Q, j > 0$ trivially hold. Thus we get $uv^iwx^iy \in Q, i \geq 0$ with $u = p, v = a, w = x = \lambda, y = q$. □

Dömösi, S. Horváth, M. Ito [1991]:

Theorem 4 *The language of all non-primitive words over a non-trivial alphabet is not context-free.*

Proof: Given the language Q over a non-trivial alphabet Σ^*, let us suppose that, contrary to our statement, $\Sigma^* \setminus Q$ is context-free. By Theorem 1, there exists a positive integer n with the following property: each word z in $\Sigma^* \setminus Q$, $|z| > n$, is of the form $uvwxy$, where $|vwx| \leq n$, $|vx| > 0$, and $uv^m wx^m y$ is in $\Sigma^* \setminus Q$, for all $m \geq 0$.

Let $a, b \in \Sigma, a \neq b$ such that $(a^{n+1}b^{n+1})^2$ is of the form $uvwxy$ with $|vwx| \leq n$, $|vx| > 0$, and $uv^m wx^m y$ is in $\Sigma^* \setminus Q$, for all $m \geq 0$. Then for $m = 0$ we have $uwy \in \{a^i b^j a^s bt \mid i, j, s, t \geq 1, (i, j) \neq (s, t)\} \subseteq Q$, contradicting $uwy \in \Sigma^* \setminus Q$. □

Stated by P. C. Fischer 1963 without proof; shown by L. H. Haines [1965] and also by S. Ginsburg and S. A. Greibach [1966] :

Theorem 5 *A language $L \subseteq \Sigma^*$ is deterministic context-free if and only if $\Sigma^* \setminus L$ is deterministic context-free.* □

Dömösi, S. Horváth, M. Ito [1991]:

Corollary 6 *Q is not deterministic context-free.*

In the next two statements we shall use two types of marked positions : "distinguished" and "excluded" positions, respectively, such that the same position may be distinguished and excluded at a time.

Ch. Bader and A. Moura [1982] :

Theorem 7 [*Bader-Moura Lemma*] *For any context-free language L, there exists a positive integer n such that for every $z \in L$, if $\delta(z)$ positions in z are "distinguished" and $\epsilon(z)$ positions are "excluded," with $\delta(z) > n^{\epsilon(z)+1}$, then there are u, v, w, x, y such that $z = uvwxy$ and*

- *(i)* *vx contains at least one distinguished position and no excluded positions;*
- *(ii)* *if r is the number of distinguished positions and s is the number of excluded positions in vwx, then $r \leq n^{s+1}$;*
- *(iii)* *for every positive integer i, $uv^i wx^i y \in L$.*

S. Horváth [1986] :

Theorem 8 [*Strong Bader-Moura Lemma*] *For any context-free language L, there exists a positive integer n depending only on L such that for every $z \in L$, if $\delta(z)$ positions in z are "distinguished" and $\epsilon(z)$ positions are "excluded," with $\delta(z) > n^{\epsilon(z)+1}$, then there are u, v, w, x, y such that $z = uvwxy$ and*

- *(i)* *either each of u, v, w or each of w, x, y contains a distinguished position and vx contains no excluded positions;*
- *(ii)* *if r is the number of distinguished positions and s is the number of excluded positions in vwx, then $r \leq n^{s+1}$;*
- *(iii)* *for every positive integer i, $uv^iwx^iy \in L$.* $\square$

Non-empty version : all of u, v, w, x, y are nonempty

Dömösi, M. Ito, M. Katsura and C. Nehaniv [1996] :

Theorem 9 (Dömösi-Ito-Katsura-Nehaniv Lemma) *For any context-free grammar G, there exists an effectively computable positive constant c depending only on G such that for any $z \in L(G)$, if $|z| \geq ce$ $(e > 0)$ and e positions of z are excluded, then z has the form $uvwxy$ where $|vx| > 0$, vx does not contain any excluded positions, and uv^iwx^iy is in $L(G)$ for all $i \geq 0$.*

Dömösi, S. Horváth, M. Ito, L. Kászonyi, and M. Katsura [1993] :

Theorem 10 *Q satisfies the condition of the Non-empty Version of Strong Bader-Moura Lemma with Bader-Moura constant $n = 5$, which is the smallest possible such constant for Q. Moreover, we can always effectively find a suitable iteration factorization (if the distinguishing-excluding condition is fulfilled in the given primitive word).*

2 Interchanging property

W. Ogden, R. J. Ross, K. Winklmann [1982] :

Theorem 11 [*Interchanging Lemma*] *for every context-free language L there exists constant $c_L > 0$ such that for all positive integers n, m with $n \geq m \geq 2$, and all subsets $H \subseteq L \cap \Sigma^n$ there exists $Z = z_1, z_2, \ldots, z_k \subseteq H$ with $k \geq \frac{|H|}{c_L \cdot n^2}$ and words $z_i, i = 1, \ldots, k$ such that*

(i)	$z_i = w_i x_i y_i, i = 1, \ldots, k$
(ii)	$\lvert w_1\rvert = \lvert w_2\rvert = \cdots = \lvert w_k\rvert$,
(iii)	$\lvert y_1\rvert = \lvert y_2\rvert = \cdots = \lvert y_k\rvert$,
(iv)	$\frac{m}{2} < \lvert x_1\rvert = \lvert x_2\rvert = \cdots = \lvert x_k\rvert \leq m$,
(v)	$w_i x_j y_i \in L \cap \Sigma^n$ *for all* $1 \leq i, j \leq k$.

We say that a language $L \subseteq \Sigma^*$ satisfies the *strengthened interchange property* if and only if there is $c > 1$ (depending only on L) such that for all $n \geq 2, i \geq 0$ and $j \geq 1$ with $j < n$ and $i + j \leq n$, and for all non-empty subsets H of $L \cap \Sigma^n$, there is $H' \subseteq H$ with the following properties: $|H'| > \frac{|H|}{c}$ and for *any two words* x and y in H', if $x = x_1x_2x_3$ and $y = y_1y_2y_3$ with $|x_1| = |y_1| = i, |x_2| = |y_2| = j$, we have $x_1y_2x_3, y_1x_2x_3 \in L$, and in this case we shortly say that H' (and also, that any pair of elements of H') is *$i - j$-interchangeable.*

We remark that the strengthened interchange property is much stronger that the Ogden, Ross and Winklmann's interchange property since in the former property we have $|H'| > \frac{|H|}{cn^2}$, and also, unlike the original Ogden, Ross and Winklmann's interchange property, in the strengthened interchange property there are no restrictions on the beginning (i) and length (j) of the middle subwords to be exchanged, other than excluding the trivial cases $j = 0$ and $j = n$.

S. Horváth [1995] :

Theorem 12 *Q satisfies the strengthened interchange property (with $c = 8$, moreover, even $c = 4$ is enough in the following three cases:*
(1) n is of the form $n = 2^k, k \geq 1$;
(2) n is odd;
(3) n is of the form $n = 2pk$ where p is an odd prime, the smallest odd prime divisor of n, $k \geq 1$, and $min\{j, n - j\} \leq \frac{n(p-1)}{2p} = (p-1)k$ (the latter condition is simply implied by $min\{j, n - j\} \leq \frac{n}{3}$.)

S. Horváth [1995] :

Theorem 13 *Q is nonlinear.*

3 Kászonyi-Katsura Theory

The Kászonyi-Katsura Theory asserts that the intersection of Q and any member of a special, infinite class of regular languages, is a context-free language.

Y. Bar-Hillel, M. Perles, and E Shamir [1961] :

Theorem 14 *L is context-free if and only if for every regular language R, $L \cap R$ is context-free.*

Definition 15 A set $F \subseteq \mathbb{N}^m$ where $\mathbb{N} = \{0, 1, \ldots\}$ and $m \geq 1$ is called a *stratified linear set* if and only if either $F = \emptyset$ or there exist $r \geq 1$ and $v_0, \ldots, v_r \in \mathbb{N}^m$ such that

(1). $F = \{ v_0 + \sum_{i=1}^{r} k_i v_i \mid k_i \geq 0 \}$

and for the vector set $P = \{ v_i \mid 1 \leq i \leq r \}$

(2). every $v \in P$ has at most two non-zero components, and

(3). there exist no natural numbers i, j, k, l, with $0 \leq i < j < k < l \leq m-1$, and no vectors $u = (u_0, \ldots, u_{m-1})$ and $x = (x_0, \ldots, x_{m-1})$ from P such that $u_i x_j u_k x_l \neq 0$.

The vector v_0 and the vector-set P appearing in (1) are often called *preperiod* and the set of *periods* of F, respectively.

Therefore, in short, E is a *stratified linear set* if it is a linear set with a stratified set of periods.

S. Ginsburg and E. H. Spanier [1964] :

Theorem 16 (Ginsburg-Spanier Theorem) *Let L be a bounded language over the alphabet Σ. Language L is context-free if and only if set*

$$E(L) = \{ (e_0, \ldots, e_{m-1}) \in \mathbb{N}^m \mid w_0^{e_0} \ldots w_{m-1}^{e_{m-1}} \in L \}, \tag{1}$$

where the words $w_0, \ldots, w_{m-1}$ are the corresponding words of L, is a finite union of stratified linear sets.

unpublished result of Kászonyi [2011] :

Theorem 17 *If* $(1/p_1 + \cdots + 1/p_k) + 1/p_1p_2 < 1$ *and* $(1/p_1 + \cdots + 1/p_k) + 1/p_3 < 1$ *then* Q_n *is context-free.*

P. Dömösi, S. Horváth, M. Ito, L. Kászonyi, and M. Katsura [1993] :

Theorem 18 *Let* $a, b \in \Sigma$, $a \neq b$, $n = p^r$ *or* $n = p^r q^s$, *where* p, q *are different prime numbers,* $r, s \geq 1$. *Let further* $L = (ab^*)^n$ *or* $L = (a^+b^+)^n$. *Then* $Q \cap L$ *is a context-free language.*

L. Kászonyi and M. Katsura [1999] :

Theorem 19 *Let* $a, b \in \Sigma$, $a \neq b$ *and* $n = p^{f_1} q^{f_2} r^{f_3}$, *where* p, q *and* r *are pairwise different prime numbers,* $f_1, f_2, f_3 \geq 1$. *Let further* $L = (ab^*)^n$. *Then* $Q \cap L$ *is a context-free language.*

4 Kászonyi's Conjecture

L. Kászonyi [1997] (?) :

Conjecture 20 *Let* $a, b \in \Sigma$, $a \neq b$ *and* n *be an arbitrary positive integer. Then* $Q \cap (ab^*)^n$ *is a context-free language.*

Maybe even a more general statement is true.

Problem: Is $L \cap Q$ context-free for every bounded language L?

Some further steps in this direction:

P. Dömösi, C. Martin-Vide, V. Mitrana [2004] :

Theorem 21 *For any slender context-free language* L, *the set* $L \cap Q$ *is also context-free.* $\square$

P. Dömösi, C. Martin-Vide, and A. Mateescu [2005] (It can also be directly derived from the results in D. Raz [1997], L. Ilie, G. Rozenberg, A. Salomaa [2000]) :

Theorem 22 *Every bounded context-free language is a finite union of non-crossing multiple paired loop languages.*

5 The proof of Theorem 21

H. J. Shyr and G. Thierrin [1977] :

Theorem 23 *Let* $i \geq 1$ *and* $uv \in \{p^i : p \in Q\}$. *Then* $vu \in \{p^i : p \in Q\}$, *too. Therefore,* $uv \in Q$ *for some* $u, v \in \Sigma^*$ *if and only if* $vu \in Q$. *In other words, the sets* $\{p^i : p \in Q\}$ $(i \geq 1)$ *are closed under cyclic permutations of words.*[2]

R. C. Lyndon and M. P. Schützenberger [1962] :

Theorem 24 *Let* $f, g \in Q, f \neq g$. *If* $fg^n \notin Q$ *then* $fg^{n+2} \in Q$ *for all* $n \geq 2$.

R. C. Lyndon and M. P. Schützenberger [1962] :

Theorem 25 *If* $u \neq \lambda$, *then there exists a unique primitive word* f *and a unique integer* $k \geq 1$ *such that* $u = f^k$. □

In this case we put $\sqrt{u} = f$ and say that f is the primitive root of u.

M. Ito, M. Katsura, H. J. Shyr and S. S. Yu [1988] :

Proposition 26 *Let* $p, q \in \Sigma^+$ *such that* $\sqrt{p} \neq \sqrt{q}$. *Then* $|pq^+ \setminus Q| \leq 1$. □

P. Dömösi and G. Horváth [2005], also (directly) from M. Ito, M. Katsura, H. J. Shyr and S. S. Yu [1988] :

Theorem 27 *Let* $f, g \in Q, f \neq g$ *and* $n \geq 1$. *If* $fg^n \notin Q$ *then* $fg^{n+k} \in Q$ *for all* $k \geq 1$.

[2]By $i = 1$ this obviously means that $uv \in Q$ for some $u, v \in \Sigma^*$ if and only if $vu \in Q$.

Almost trivial (P. Dömösi, C. Martin-Vide, V. Mitrana [2004]) :

Proposition 28 *Let* $ac, b \in \Sigma^+$ *such that* $\sqrt{ca} \neq \sqrt{b}$. *Then* $|ab^+c \setminus Q| \leq 1$.

Proof: Using Theorem 23, it is enough to prove that $|cab^+ \setminus Q| \leq 1$ whenever $ac, b \in \Sigma^+$ such that $\sqrt{ca} \neq \sqrt{b}$. But this is a direct consequence of Proposition 26. □

M. Latteux and G. Thierrin [1983] and later, independently, by L. Ilie [1994] and D. Raz [1997] :

Theorem 29 *Every slender context-free language is a finite disjoint union of paired loop languages (DUPL in short).* □

P. Dömösi, C. Martin-Vide, V. Mitrana [2004] :

Proposition 30 *Let* $ace, b, d \in \Sigma^+$ *with* $|\{k : \sqrt{eab^kc} = \sqrt{d}\}| = \infty$. *Then* $\{ab^ncd^ne : n \geq 1\} \cap Q = \emptyset$.

Proof: Case 1. $eac = \lambda$. Then, by $|\{k : \sqrt{eab^kc} = \sqrt{d}\}| = \infty$, there exist infinite-many $k \geq 1$ with $\sqrt{b^k} = \sqrt{d}$. On the other hand, for every $k \geq 1$, we have $\sqrt{b^k} = \sqrt{d}$ if and only if $\sqrt{b} = \sqrt{d}$. But this implies $b^kd^k \notin Q, k \geq 1$.

Case 2. $eac \neq \lambda$. First we prove that $\sqrt{cea} \neq \sqrt{b}$ is impossible. Indeed, assume $\sqrt{cea} \neq \sqrt{b}$. If $cea \notin Q$, then by Theorem 24, $ceab^n \in Q, n \geq 2$. If $cea \in Q$, then by Theorem 27, $ceab^n \in Q, n \geq 3$. Therefore, by Theorem 23, $eab^nc \in Q, n \geq 3$. But then for every $s, t \geq 3$, we obtain $\sqrt{eab^sc} = \sqrt{eab^tc}$ if and only if $s = t$. Therefore, if $\sqrt{eab^kc} = \sqrt{d}$ then $\sqrt{eab^{k+\ell}c} \neq \sqrt{d}, \ell \geq 1$. But then $|\{k : \sqrt{eab^kc} = \sqrt{d}\}| < \infty$, a contradiction. Thus, we have $\sqrt{cea} = \sqrt{b}$ (with $eac \neq \lambda$). But then $\sqrt{eab^sc} = \sqrt{eab^tc}, s, t \geq 1$.

On the other hand, by $|\{k : \sqrt{eab^kc} = \sqrt{d}\}| = \infty$, there exist infinite-many $k \geq 1$ having $\sqrt{eab^kc} = \sqrt{d}$. Hence, using $\sqrt{eab^sc} = \sqrt{eab^tc}, s, t \geq 1$, we obtain $\sqrt{eab^kc} = \sqrt{d}, k \geq 1$. Thus, we get $\{ab^ncd^ne : n \geq 1\} \cap Q = \emptyset$ as we stated. □

P. Dömösi, C. Martin-Vide, V. Mitrana [2004] :

Proposition 31 *Let* $ace, b, d \in \Sigma^+$ *with* $|\{k : \sqrt{eab^k c} = \sqrt{d}\}| < \infty$. *Then* $|\{ab^n cd^n e : n \geq 0\} \setminus Q| < \infty$.

Proof: Case 1. $d \neq (eab^i c)^j, i \geq 0, j \geq 1$.

Observe that we have either $b = (cea)^s$ for some $s \geq 1$ or there exists an $\ell \geq 1$ such that $eab^n c \in Q$ for all $n \geq \ell$. Indeed, assume $b \neq (cea)^s, s \geq 1$. If $eac \in Q$ then we can apply Proposition 28. Otherwise, by Theorem 23, $eac, cea \in \{q^i : q \in Q\}$ for some $i \geq 2$. Then, by Theorem 24, $ceab^n \in Q, n \geq 2$. Considering Theorem 23, this implies $eab^n c \in Q, n \geq 2$.

Assume $b = (cea)^s$ for some $s \geq 1$. Having $d \neq (eab^i c)^j, i, j \geq 1$, we may apply Theorem 24 such that $eab^n cd^n \in Q, n \geq 2$. Thus, by Theorem 23, $ab^n cd^n e \in Q, n \geq 2$.

It remains to study the case when there exists an $\ell \geq 1$ such that $eab^n c \in Q$ for all $n \geq \ell$. Thus, applying Theorem 23, $ab^n ce \in Q$ for all $n \geq \ell$. But then, considering Proposition 28 and assuming $d \neq (eab^i c)^j, i \geq 0, j \geq 1$, there exists a $k \geq \ell$ such that $ab^n cd^n e \in Q$.

Case 2. $d = (eab^i c)^j$ for some $i \geq 0, j \geq 1$.

Then consider $ea'b^n c'd^n, n \geq 0$ instead of $eab^n cd^n, n \geq 1$ such that $a' = ab^{i+1}$ and $c' = cd^{i+1}$. Obviously, $d \neq (ea'b^s c')^t$. Thus we may apply the previous case such that $a'b^n c'd^n e \in Q$ whenever $n \geq k$ for an appropriate $k \geq 1$. But then $ab^n cd^n e \in Q$ whenever $n \geq i + k + 1$. $\square$

P. Dömösi, C. Martin-Vide, V. Mitrana [2004] :

Theorem 32 *Let L be a DUPL such that $L = \bigcup_{i=1}^{k}\{u_i v_i^n w_i x_i^n y_i : n \geq 0\}$ for some positive k and words $u_i, v_i, w_i, x_i, y_i, 1 \leq i \leq k$ with $\{u_i v_i^n w_i x_i^n y_i : n \geq 0\} \cap \{u_j v_j^n w_j x_j^n y_j : n \geq 0\} = \emptyset, 1 \leq i < j \leq k$. Then $L \cap Q$ is also a DUPL such that $L = \bigcup_{i=1}^{2k} L_i$ with $L_i \cap L_j = \emptyset, 1 \leq i < j \leq 2k$, where for every $1 \leq i \leq k$, either $L_i = \emptyset$ with $L_{i+k} \in \{\{u_i w_i y_i\}, \emptyset\}$, or $L_i = \{u_i v_i^{\ell_i + n} w_i x_i^{\ell_i + n} y_i : n \geq 0\}, \ell_i \geq 0$, $L_{i+k} = \emptyset$ if $\ell_i = 0$, $L_{i+k} \subseteq \bigcup_{j=0}^{\ell_i - 1} u_i v_i^n w_i x_i^n y_i$ if $\ell_i > 0$.*

Proof: Proof: It is enough to prove that for every $1 \leq i \leq k$, $\{u_i v_i^n w_i x_i^n y_i : n \geq 0\} \cap Q = \{u_i v_i^{\ell_i + n} w_i x_i^{\ell_i + n} y_i : n \geq 0\} \cup L_{i+k}, \ell_i \geq 0$ such that $L_{i+k} = \emptyset$ if $\ell_i = 0$ and $L_{i+k} \subseteq \bigcup_{j=0}^{\ell_i - 1} u_i v_i^n w_i x_i^n y_i$ if $\ell_i > 0$.

If $u_i = v_i = \lambda$ holds for some $1 \leq i \leq k$, then $\{u_i v_i^n w_i x_i^n y_i : n \geq 0\}$ obviously has this property. If $u_i = \lambda$ and $v_i \neq \lambda$, or symmetrically, if $u_i \neq \lambda$ and $v_i = \lambda$, then we get the above property applying Proposition 28.

Let $u_i, v_i \neq \lambda$ for some $1 \leq i \leq k$ and suppose $|\{k : \sqrt{y_i u_i v_i^k w_i} = \sqrt{x_i}\}| = \infty$. Then we may apply Proposition 30.

Now let $u_i, v_i \neq \lambda$ for some $1 \leq i \leq k$ and suppose $|\{k : \sqrt{y_i u_i v_i^k w_i} = \sqrt{x_i}\}| < \infty$. Then we can use Proposition 31. □

By Theorem 32, we know that for every slender context-free language L, the language $L \setminus Q$ is a DUPL language. Thus we have the next statement.

P. Dömösi, C. Martin-Vide, V. Mitrana [2004] :

Corollary 33 *For any slender context-free language L, the set $L \cap Q$ is also context-free.* □

References

[1] Bader, Ch.; Moura, A.: A generalization of Ogden's lemma. Journ. of ACM, **29** (1982), 404–407.

[2] Bar-Hillel, Y.; Perles, M.; Shamir, E.: On formal properties of simple phrase structure grammars. Zeitschrift für Phonetik, Sprachwuissenschaft, und Kommunikationsforschung, **14** (1961), 143–177.

[3] Dömösi, P.; Horváth, G.: On products of primitive words. In: Ésik, Z. & Fülöp, Z., eds., Proc. 11th Int. Conf. on Automata and Formal languages (AFL'05), Dobogókő, Univ. Szeged, Inst. Inform., Szeged (2005), 112–121.

[4] Dömösi, P.; Horváth, G.: Alternative Proof of the Lyndon- Schẗzenberger Theorem. Theoret. Comput. Sci. **366** No.3, (2006), 194–198.

[5] Dömösi, P.; Horváth, S.; Ito, M.: On the connection between formal languages and primitive words. In: Proc. First Session on Scientific Communication, Univ. Of Oradea, Oradea, Romania, 6-8 June, 1991, Anns. Univ. Of Oradea (Analele Univ. din Oradea), Fasc. Mat., 1991, 59–67.

[6] Dömösi, P.; Horváth, S.; Ito, M.: Formal languages and primitive words. Publ. Math. Debrecen 42 (1993), no. 3-4, 315–321.

[7] Dömösi, P.; Horváth, S.; Ito, M.; Kászonyi, L.; Katsura, M.: Formal languages consisting of primitive words. In: Ésik, Z., ed., Fundamentals of computation theory (Szeged, 1993), 194–203, Lecture Notes in Comput. Sci., **710**, Springer, Berlin, 1993.

[8] Dömösi, P.; Horváth, S.; Ito, M.; Kászonyi, L.; Katsura, M.: Some combinatorial properties of words, and the Chomsky-hierarchy. In: Ito, M. and Jürgensen, H., eds., Words, languages and combinatorics, II (Kyoto, 1992), 105–123, World Sci. Publishing, River Edge, NJ, 1994.

[9] Dömösi, P.; Ito, M.; Katsura, M.; Nehaniv, C., L.: A new pumping property of context-free languages. In: Bridges, D., S., Calude, S., Gibbons, J., Reeves, S., and Witten, I., H., eds., Combinatorics, complexity, & logic (Auckland, 1996), 187–193, Springer Ser. Discrete Math. Theor. Comput. Sci., Springer, Singapore, 1997.

[10] Dömösi, P.; Martin-Vide, C.; Mateescu, S.: On Polyslender Context-Free Languages. Publicationes Math., Debrecen **66** (2005), 1–15.

[11] Dömösi, P.; Martin-Vide, C.; Mitrana, V. : Remarks on sublanguages consisting of primitive words of slender regular and context-free languages. In: Păun, Gh. ed., Theory is Forever, 60–67, Lecture Notes in Comput. Sci., **3113**, Springer, Berlin, 2004.

o. 3-4,

[12] Ginsburg, S. ; Greibach, S. A. : Deterministic context-free languages. Inform. and Contr., **9** (1966), 620–648.

[13] Ginsburg, S. ; Spanier, E. H.: Bounded ALGOL-like languages. Trans. Am. Math. Soc. **113** (1964), 333–368.

[14] Haines, L. H. : Generation and Recognition of Formal Languages. doctoral dissertation, Mass. Inst. of Techn., Cambridge, Mass., June, 1965.

[15] Fischer, P. C. : On computability by certain classes of restricted Turing machines. Proc. Fourth. Ann. Symp. on Switching Circuit Theory and Logical Design, Chicago, 28-30 Oct., 1963, 23–31.

[16] Horváth, S.: The family of languages satisfying Bar-Hillel's lemma. RAIRO Inform. Théor. 12 (1978), no. 3, 193–199.

[17] Horváth, S.: A comparison of iteration conditions on formal languages. In: Demetrovics, J., Katona, G., Salomaa, A., eds., Algebra, combinatorics and logic in computer science, Vol. I, II (Györ, 1983), 453–463, Colloq. Math. Soc. János Bolyai, 42, North-Holland, Amsterdam, 1986.

[18] Horváth, S.: Strong interchangeability and nonlinearity of primitive words. In: Nijholt, A. Scollo, G, Steetskamp, R., eds., Proc. Works. AMilP'95 (Algebraic methods in language Processing, 1995) Univ. of Twente, Enschede, the Netherlands, 6-8 Dec., 1995, Univ. Twente, Enschede, 1995, 173–178.

[19] Ilie, L.: On a conjecture about slender context-free languages. Theoret. Comput. Sci., 132 (1994), 427–434.

[20] Ilie, L.; Rozenberg, G.; Salomaa, A.: A characterization of poly-slender context-free languages. RAIRO, Theoret. Inform. Appl., 34 (2000), no. 1, 77–86.

[21] Ito, M.; Katsura, S.: Context-free languages consisting of non-primitive words. Semigroup Forum 37 (1988), 45–52.

[22] Ito, M.; Katsura, M.; Shyr, H. J.; Yu, S. S.: Automata accepting primitive words. Semigroup Forum 37 (1988), no. 1, 45–52.

[23] Kászonyi, L.: The Language $Q \cap (ab^*)^n$ with some restrictions on the prime divisors of n. Manuscript (2011), 1–4.

[24] Kászonyi, L.; Katsura, M.: On the context-freeness of a class of primitive words. Publ. Math. Debrecen 51 (1997), no. 1-2, 1–11.

[25] Kászonyi, L.; Katsura, M.: Some new results on the context-freeness of languages $Q \cap (ab^*)^n$. In: Ádám, A., Dömösi, eds., Automata and formal languages, VIII (Salgtarján, 1996). Publ. Math. Debrecen 54 (1999), suppl., 885–894.

[26] Latteux, M; Thierrin, G.: Semidiscrete context-free languages. Internat. J. Comput. Math. 14 (1983), 3–18.

[27] Leupold, P.: On some properties of context-free languages related to primitive words. In: Harju, T., Karhumäki, J., eds., Proceedings of WORDS'03, 282–291, TUCS Gen. Publ., 27, Turku Cent. Comput. Sci., Turku, 2003.

[28] Leupold, P.: Primitive Words Are Unavoidable for Context-Free Languages. Dediu, Adrian-Horia (ed.) et al., Language and automata theory and applications. 4th international conference, LATA 2010, Trier, Germany, May 24–28, 2010. Proceedings. Berlin, Springer. Lecture Notes in Computer Science 6031, 403–413 (2010).

[29] Lyndon, R. C.; Schützenberger, M. P.: The equation $a^m = b^n c^p$ in a free group. Michigan Math. J., 9 (1962), 289–298.

[30] Ogden, W.; Ross, R. J.; Winklmann, K.: An "Interchange Lemma" for context-free languages. SIAM J. Comp. 14 (1985), 410–415.

[31] Raz, D.: Length considerations in context-free languages. Theoret. Comput. Sci., **183** (1997), 21–32.

[32] Reis, C. M.; Shyr, H. J.: Some properties of disjunctive languages on a free monoid. Inform. and Contr., **37** (1978), 334–344.

[33] Shyr, H. J.; Thierrin, G.: Disjunctive languages and codes. In: Karpiński, M., ed., Proc. Conf. FCT'77, vol 56, Springer-Verlag, 1977, 171–176.